INTELLECTUAL ASSESSMENT IN PRIMITIVE SOCIETIES, WITH A PRELIMINARY REPORT OF A STUDY OF THE EFFECTS OF EARLY IODINE SUPPLEMENTATION ON INTELLIGENCE

Frederick L. Trowbridge

Center for Disease Control

Atlanta, Georgia

Mental deficiency is perhaps the most tragic aspect of endemic cretinism. Yet, little has been done to define the nature and extent of intellectual deficiency in areas where endemic cretinism is widely seen. The mental deficiency of a typical cretin is obvious, but it is not known whether iodine deficiency or other factors which may lead to endemic cretinism can also cause variable and less obvious intellectual deficits in individuals who are not cretins. Another important question relates to time factors in the effects of iodine deficiency on intellectual development. It has been presumed that the most critical period is in early development. However, one reported study suggested that iodine supplementation as late as four to eight years of age might have some effect on measured intelligence (1). This finding is theoretically difficult to interpret, but indicates the need for further study of this important problem.

MEANS OF INTELLECTUAL ASSESSMENT

Perhaps the reason that so little has been done in the assessment of intelligence in areas of endemic cretinism is that the tools for assessing intelligence in primitive societies are poorly developed. Intelligence is an elusive entity, and intelligence tests under any conditions are no more than estimates of mental ability. When tests are applied outside of the culture in which they were developed, additional problems arise. Most fundamentally, these stem from the different way in which reality is perceived, the different manner of interpreting problems, and the different modes of thought which characterize each culture. Because of these differences tests designed to assess intelligence as defined by one culture may be invalid in another.

Many extrinsic factors can affect test results when tests are inadequately adapted to the local culture. The novelty and strangeness of the test situation may be frightening or upsetting. Motivation to perform well can be culturally determined and highly variable. Instructions may be misunderstood. Materials may seem foreign and meaningless or may hold unintended meanings or associations leading to incorrect evaluation of responses.

Broad biological and social factors also affect test scores. Poor nutrition, infection, crowding, lack of intellectual stimulation, inferior language experience, and other environmental factors can influence the results of any intelligence assessment. In underdeveloped areas such conditions may be starkly evident, but their specific effect on intellectual development is extremely difficult to quantify. This reality confounds attempts to identify or test specific causal factors as they relate to intellectual development.

The many hazards encountered in attempting intellectual assessment in different cultures has emphasized the need for tests less affected by cultural "bias." Important progress has been made in this direction. In general the attempt has been to minimize language-dependent items and to eliminate materials which would be familiar in one culture but unfamiliar in another. There are many examples of these so-called "culture-fair" tests. Among the more widely used are tests of "inter-sensory modalities" developed by Birch and Lefford (2, 3). This testing technique evaluates the integration of perception through visual, kinesthetic, and "haptic" modalities. The haptic modality refers to the perception of the shape of an object which the subject feels but cannot see. In testing kinesthetic perception the subject's hand holds a stylus which is guided around the outline of an object to be identified. Application of inter-sensory testing techniques has been made by Cravioto in assessing the interaction of nutrition and mental development in Mexican school-age children (4). Klein has adapted inter-sensory techniques for application to a rural Guatemalan culture (5).

Other tests which utilize materials that can be administered cross-culturally are Raven's Progressive Matrices (6) and Kohs Blocks. These tests present non-verbal figurative problems in the form of objects or designs which must be correctly aligned or rearranged to complete a prescribed pattern. There are numerous other tests of this type which have analogous format. An adaption of Kohs Block Designs has been made by Ord for use in New Guinea with subjects who have had very little exposure to Western culture (7).

Despite the important advantages of tests which avoid culture-bound features, the most fundamental problems of cross-cultural testing are not eliminated. All persons live within a cultural context, and responses to any testing situation will be molded by that culture. The so-called "culture-fair" tests may reliably demonstrate individual differences within a culture, but the significance of what is measured may vary from one culture to another.

When testing is to be done within a culture rather than between different cultures it may be practical to aim not at developing a culture-free test, but rather to adapt a standard test or develop new test procedures within that cultural context. In either case it is obvious that materials, instructions and test conditions must be suited to the local environment. Tests must also be standardized for a particular culture by the criteria of reliability and validity. Reliability is a measure of within-test and the test-to-test consistency. Items which are assigned a given level of difficulty must be passed or failed by appropriate percentages of children of various ages. The test should not be sensitive to extrinsic factors in the conditions of testing, so that tests given at different times will be comparable.

Another obvious criteria which must be applied is that of validity. For a test to be valid within a culture there must be a relationship between test results and external correlates of intelligence such as school performance or occupational achievement.

The limitations of whatever testing method is used must be recognized. When properly applied, intelligence tests can be used to measure differences in mental ability within a culture. It is always far more hazardous to attempt to compare the results obtained in one culture with those obtained in another.

EVALUATION OF INTELLIGENCE AFTER IODINE SUPPLEMENTATION: REPORT OF A STUDY

Having discussed an approach to intellectual assessment in primitive societies, a specific application of intelligence testing in an endemic goiter area is now presented. This study involved the mental assessment of children after iodine supplementation, comparing children who were exposed to supplemented levels of iodine at various times in gestation or infancy with non-supplemented control groups.

METHOD

Two populations of rural Ecuadorean children were tested. One population was drawn from a village in which iodine supplementation was given in March, 1966, by depot intramuscular injection. Supplementation was given to approximately 90 percent of individuals in the village, from newborns to aged persons. The control population was drawn from a nearby village similar in social, cultural, and economic makeup, but in which no iodine supplementation was given. Details of this iodine supplementation program and of ongoing studies in the same villages have been reported elsewhere (8).

In each village the children were divided into three groups. Group I consisted of children born in the nine-month period before the iodine was given. In the test village this group received the iodine injection some time between birth and nine months of age. Group II consisted of children born in the nine-month period immediately after iodinization. In the test village these children were born of mothers receiving the iodine injection during the gestational period. They were, therefore, exposed to supplemented levels of iodine for from less than one month to the full nine months of gestation. Group III consisted of children born from nine to 18 months after iodinization. In the test village the mothers had iodine supplementation prior to their conception. Therefore, these children experienced supplemented levels of iodine during their entire gestational period.

The ages of the children at the time of testing varied from a mean age of five years four months for Group I to a mean age of three years 10 months for Group III. Exact ages for the purpose of scoring were determined primarily from baptismal records kept by the parish priest.

In the test village all surviving children born within the dates specified for each group and who could be found and identified were tested. A slightly larger number of children in each group from the other village were tested to serve as controls. They were selected to give a roughly even distribution throughout the age range of each group. The children called to be tested each day were of mixed ages representing a sampling of all groups, and the group identification of the child was unknown to the physician administering the test. Also, testing was changed from one village to the other every two to four days to prevent any inadvertent changes in testing conditions or procedure from affecting relative performance.

The test used was an adaptation of the Stanford-Binet Intelligence Scale. The tests for ages two through eight years were translated into Spanish, employing words and phrasing judged to be most easily understood by these rural Ecuadorean children. The translation was made by an Ecuadorean physician who had lived in the test village for two years and had a thorough knowledge of the terminology and manner of speaking used by the people. The test was administered according to this translation by a North American physician with a good working knowledge of Spanish. To insure that the presentation was clear and that responses were interpreted correctly, an Ecuadorean medical student observed all tests and participated when necessary. The simplicity of the terminology used in the questions for the age levels tested was such that the children had no difficulty understanding the questions or instructions. The guidelines for giving the tests were adhered to closely. The manner of scoring also followed the standard form. New materials were substituted in a number of questions because the standard Stanford-Binet items were obviously not pertinent to the rural South American culture. Substituted material was chosen to maintain the intent of the original item. A listing and description of all substitutions and modifications is available from the author.

Although the children tested were all at least three and one half years of age, some of them failed to complete successfully all of the Stanford-Binet at the two year level. Therefore, in some cases items from the Catell Infant Intelligence Scale were used to establish a baseline at which all items were completed successfully. Once established, this baseline was used for scoring purposes in the same manner as a baseline established with the Stanford-Binet items.

The tests were carried out in a single room in each village set up for the purpose. The examiner sat across a small table from the child. To one side, behind a desk, sat an assistant who recorded the responses to each question. The Ecuadorean medical student sat slightly behind the child in order to be able to hear and observe all responses. The children in these villages were often shy, and it was found important to have the mother of the child in the room to provide a reassuring presence. The mother was seated to the rear of the child so that she could not give visible clues to the child. No verbal assistance from the mother was permitted.

Several minutes were spent initially playing "catch" with a plastic ball or playing with a doll to give the child a chance to accomodate to the new surroundings. Subsequently, testing was begun with playful items at the three year level; stringing beads on a shoestring and building a block tower and bridge. The picture vocabulary was given next since performance on this item served as an indicator of the child's verbal baseline. Testing was then continued at the level at which the picture vocabulary was successfully passed. As defined in the Stanford-Binet method testing was continued from the baseline where all items were passed, to the ceiling level at which all items were failed. The child's mental age was then calculated by adding up credit for the items completed successfully between these two levels. An IQ was then determined from the standard tables in the Stanford-Binet manual.

After testing was completed results were excluded if circumstances were encountered which would invalidate test results. Exclusions were made for the following reasons:

1. If the child was uncooperative and resistant to the testing situation to the point of being unscorable.
2. If the child was recognized as mentally deficient by the parents and others in the village.
3. If the child displayed behavior abnormalities such as excessive nervousness or hyperactivity.
4. If the score was so strikingly low as to suggest that the child was an unrecognized deficient or that he grossly failed to understand what was required by the test.

A pilot study was conducted to standardize instructions, materials and procedures. Twenty-two children from the control village were tested. These pilot-phase children were excluded from participation in the actual study.

RESULTS

A total of 125 children were tested (Table 1), 59 in the test village (Tocachi) and 66 in the control village (La Esperanza). After testing, some scores were excluded for the reasons outlined above. A total of 14 exclusions were made (Table 2). These exclusions were distributed throughout the test groups. After exclusions, the remaining total of 111 children were included in final tabulations, 51 from Tocachi and 60 from La Esperanza (Table 3).

Table 1. Children tested

Group	Tocachi	La Esperanza	Total
I	16	19	35
II	21	23	44
III	22	24	46
Total	59	66	125

Table 2. Exclusion of scores

	Tocachi			La Esperanza			
Reason:	I	II	III	I	II	III	Total
Unscorable	0	2	1	0	0	2	5
Mental Deficient	0	0	2	0	0	1	3
Behavioral Abnormality	1	0	0	0	2	0	3
Score Below 50	2	0	0	1	0	0	3
Total	3	2	3	1	2	3	14

The number of children in each group varied from 13 to 21. There was also some variation in the proportion of male and female subjects in each group. However, the total number of male and female subjects tested was not greatly different.

Social data tabulations indicated some differences between the villages (Table 4). In Tocachi there was a higher reported literacy rate and a higher percentage of fathers who worked as farmers. A lower percentage of families in Tocachi reported cretins among their close relatives.

Anthropometric data were generally comparable between the two villages. However, children in Group I in Tocachi were generally shorter and weighed less than their counterparts in Group I in La Esperanza.

Table 3. Composition of test groups - after exclusions

	Tocachi			La Esperanza			Total		
Group	M	F	Total	M	F	Total	M	F	Total
I	10	3	13	11	7	18	21	10	31
II	11	8	19	9	12	21	20	20	40
III	6	13	19	7	14	21	13.	27	40
Total	27	24	51	27	33	60	54	57	111

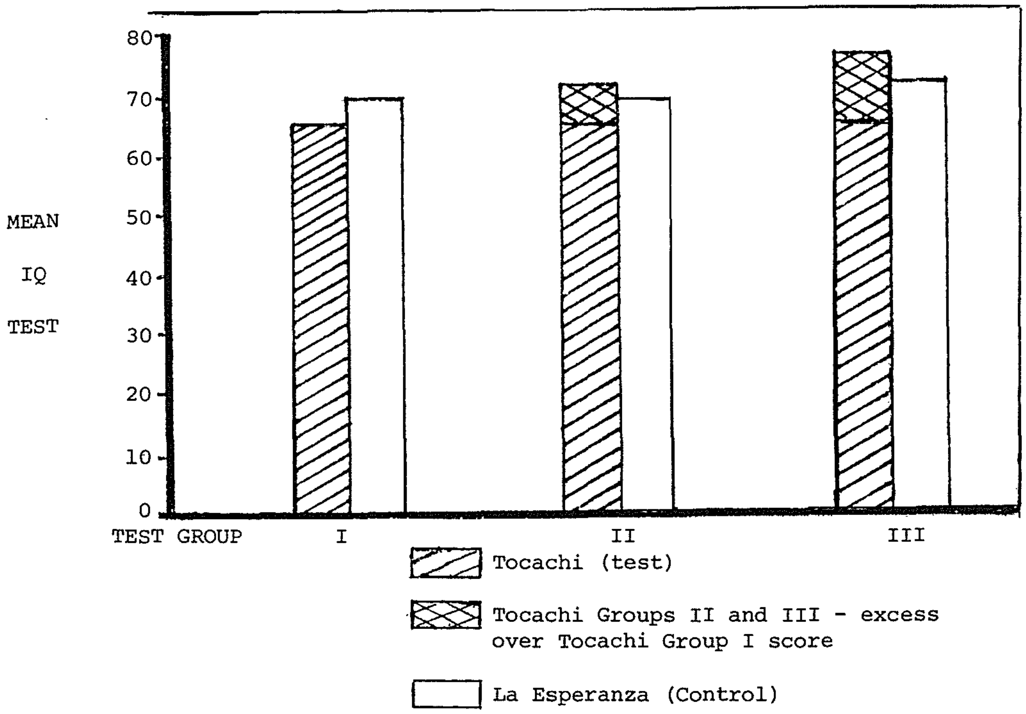
Table 4. Social data

Literacy	% Tocachi	% La Esperanza
Mother	56.3	36.4
Father	86.4	77.4
Both	42.6	33.3
Neither	4.3	21.2
Ownership		
House	60.4	68.1
Land	39.1	45.5
Radio	45.3	48.4
Father's Occupation		
Farmer	95.0	72.2
Other	5.0	27.8
Cretins Among Relatives		
	33.3	46.0

Mean IQ scores for all groups showed no significant difference between the test and control villages (Table 5). Also, within the control village there was no significant difference between the mean IQ score of the three groups. However, within the test village there was a progressive increase in mean IQ score from Group I to Group III (Table 6). Children in Group III in Tocachi, whose mothers were iodized before conception, had a mean IQ score which was significantly greater than the mean IQ score of children in Group I, who received iodine supplementation after birth ($p=.005$). Similarly, Group II children, whose mothers received iodine during the gestational period, showed a mean IQ score which was significantly higher than that of Group I in the same village ($p=.05$). Mean IQ scores for all groups are presented in Table 5. The excess in mean IQ scores of Group II and Group III is indicated. The difference between the mean IQ values of Group II and Group III in Tocachi was not significant.

Within the test groups in Tocachi there were fairly wide but not statistically significant differences between the mean IQ scores of male and female subjects. Comparing female scores within Tocachi showed a striking progression from a mean score of 61 for females in Group I, to 68 in Group II, to 78 in Group III. The difference between Group I and Group II and between Group II and Group III were statistically significant. Mean IQ scores for males in Tocachi showed a significant difference between a mean score of

Table 5. Test Results



66 in Group I and a mean score of 75 in Group II. A mean score for group III of 73 showed a slight trend in the opposite direction from the rest of the data.

Table 6. Mean I.Q. Scores

	Toacachi			La Esperanza		
GROUP	M	F	TOTAL	M	F	TOTAL
I	66	61	65.2	69	70	69.9
II	75	68	72.3	66	70	69.0
III	73	78	76.8	72	72	72.4
TOTAL			72.2			70.4

Inclusion of scores which were excluded for the reasons outlined did not substantially affect the trend or the significant implications of the data.

DISCUSSION

The most significant implications of the data relate to the mean IQ scores of Group I in Tocachi. Because this is a critical group it is important to outline some atypical characteristics which it displayed. It was the smallest group and contained an unequal sex distribution of 10 males and 3 females. As a group its height and weight values were smaller than its control group in La Esperanza. The difference stemmed predominantly from the small stature and low weight of the 3 females in this group. The irregularities in the composition of this group are interesting to note but do not negate the implication of the low mean IQ score of the group as a whole.

The mean IQ score of Tocachi Group I was distinctly although not significantly lower than the corresponding group in the control village. This lower IQ value may reflect a real difference in IQ between the two villages. It is possible that Tocachi's basal level of measurable intelligence was lower than that of La Esperanza prior to iodine supplementation, and that the improvement seen in test scores in Group II and Group III children resulted from the provision of iodine at a critical stage in gestation. Unfortunately, no pretest of intelligence in the two villages was performed. Therefore this possibility can only remain in the realm of conjecture.

The six children whose scores were excluded because of behavioral abnormalities or because of excessively low scores form an interesting group which deserves further comment. As a group these children were below the mean for their age for height, weight, and head circumference. Some were strikingly small for their age and might have been children with severe nutritional inadequacy. Perhaps of greatest significance are the extremely low scores in children who were not known mental deficients. Some children displayed behavioral abnormalities which might explain their poor score. However, other children seemed to participate well in the testing but still scored in the same range as children who were recognized as mental deficients. These children may have been unrecognized mentally retarded children.

Further statistical work is required to evaluate the within-test and the test-to-test reliability of the modified Stanford-Binet Scale used. Also, it would be valuable to further validate the tests through comparison of scores with the results of other testing and with other criteria such as future school performance.

Longitudinal studies of development using Gesell Scales have been carried out in both villages since iodine supplementation was given. Comparison of results from the Gesell Scales with those from the present study on an individual and group basis would be a valuable validation tool.

It is important to note that subjects recognized as mentally deficient did score at the extreme low end of the scale. At the opposite end of the scale, the four-year-old daughter of an Ecuadorean physician who was tested but was not part of the study had an IQ score of 125. This correspondence of test scores at both extremes with anticipated performance levels provides some rough validation of the testing.

Despite all attempts to adapt the Stanford-Binet to the local culture and to apply the test consistently, the study has obvious weaknesses. The test used has significant cultural bias. The reliability and validity of the test has not been demonstrated. While the implications of the data may be significant, it must be stressed that they are preliminary in nature and are based on data which has important qualifications and limitations.

SUMMARY

The factors which cause severe mental deficiency in cretins may also cause variable and more subtle mental deficits in individuals from the same environmental and genetic background. The extent of mental deficiency in areas of endemic cretinism and the effects of iodine prophylaxis on intellectual development have been little studied, in large part because of methodological difficulties in making valid intellectual assessments in primitive societies.

An approach to intellectual assessment is outlined, and a study is reported on the effects of iodized oil prophylaxis at different stages of gestation and early infancy. Results are preliminary and qualified but suggest that iodine prophylaxis of mothers before conception or during gestation may improve subsequent intelligence test performance of their children when compared with the performance of children who were not exposed to iodine supplementation until after birth.

Validity of testing increases with age. It would be advantageous to retest the same groups as the children grow older. This would provide data for comparison with the findings of the present study.

Comparison of scores with the results of ongoing longitudinal studies in the same villages would provide additional validation.

Further delineation of the interaction of iodine prophylaxis and mental development is required. It is proposed that future programs of iodine prophylaxis in areas of endemic goiter should include intellectual assessment of children in both preliminary and follow-up evaluation. The need to establish pre-treatment norms for intelligence test performance is stressed.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the assistance of Dr. John Stanbury and Dr. Philip Dodge in formulating the study, and the collaboration of Dr. Rodrigo Fierro-Benitez and Dr. Ignacio Ramirez, in adapting the study for use in Ecuador. Mrs. Jane Trowbridge and Mr. Jose Suarez assisted in the testing procedure. Miss Mary Shepherd assisted by typing the manuscript.

This study was supported in part by U. S. Public Health Service Training Grant No. AM05618.

DISCUSSION BY PARTICIPANTS

PRETELL: We have been doing similar studies in Peru. We have tested by two different groups of psychologists and by different tests, and there was good correlation among the groups.

BALÁZS: I would like to point out that iodine supplementation was left out during a crucial period of the development of the CNS, that is during the immediate postnatal period.

STANBURY: Don't you think that there was supplementation in the sense that there would be enough transplacental transfer to build up quite adequate stores in the newborn?

PRETELL: We were able to test urinary excretion of iodine from control children and children born to iodine-injected mothers about two years after birth. In children born to non-treated mothers the iodine excretion for 24 hours was about 20 μg. In children born to treated mothers, it was about 67 μg. That probably was due to transfer during gestation, but also after birth by lactation. We also measured iodine excretion in milk and found it was much higher in the treated mothers.

REFERENCES

1. Dodge, P.R. *et al*.: Effect on intelligence of iodine in oil administered to young children. A preliminary report. In: *Endemic Goiter*. Report of the meeting of the PAHO Scientific Group on Research in Endemic Goiter held in Mexico, 1968 (J.B. Stanbury, ed.). PAHO Scientific Publication No. 193, Washington, 1969. pp 378-380.

2. Birch, H.G. and Lefford, A.: Intersensory development in children. *Monogr. Soc. Res. Child Dev.* Serial No. 89:1-48, 1963.

3. Birch, H.G. and Lefford, A.: Visual differentiation. Intersensory integration and voluntary motor control. *Monogr. Soc. Res. Child Dev.* Serial No. 110:1-77, 1967.

4. Cravioto, J.; Delicarde, E.R. and Birch, H.G.: Nutrition, growth and neurointegrative development: An experimental and ecological study. *Pediatrics*. 38:319-372, 1966.

5. Klein, R.E. Unpublished. Address requests for information to Dr. Robert Klein, Director, Human Development Division, Instituto de Nutricion de Centro America y Panama, Guatemala City, Guatemala.

6. Birch, H.R.: Raven's progressive matrices. *J. Genet. Psychol.* 93:199-228, 1958.

7. Breshuvel, S., ed.: Methods for the measurement of psychological performance. *International Biological Programme*, Blackwell Scientific Publications, Oxford, 1969. pp. 87-93

8. Fierro-Benítez, R.; Ramirez, I.; Estrella, E.; Jaramillo, C.; Diaz, C. and Urresta, J.: Iodized oil in the prevention of endemic goiter and associated defects in the Andean region of Ecuador. In: *Endemic Goiter*. Report of the meeting of the PAHO Scientific Group on Research in Endemic Goiter held in Mexico, 1968 (J.B. Stanbury, ed.). PAHO Scientific Publication No. 193, Washington, 1969. pp. 306-394.

A BIBLIOGRAPHY - INTELLECTUAL ASSESSMENT IN PRIMITIVE SOCIETIES

Anastasi, A.: *Psychological Testing*. Macmillan, New York, 1968.

Birch, H.G. and Lefford, A.: Intersensory development in children. *Monogr. Soc. Res. Child Dev.* Serial No. 89:1-48, 1963.

Birch, H.G. and Lefford, A.: Visual differentiation. Intersensory integration and voluntary motor control. *Monogr. Soc. Res. Child Dev.* Serial No. 110:1-77, 1967.

Birch, H.G.: Raven's progressive matrices. *J. Genet. Psychol.* 93:199-228, 1958.

Breshuvel, S., ed.: Methods for the measurement of psychological performance. In: *International Biological Programme*. Blackwell Scientific Publications, Oxford, 1969. pp 87-93.

Cravioto, J.; Delicarde, E.R. and Birch, H.G.: Nutrition, growth and neurointegrative development: An experimental and Ecologic Study. *Pediatrics*. 38:319-372, 1966.

Dodge, R.R. *et al*.: Effect on intelligence of iodine in oil administered to young children. A preliminary report. In: *Endemic Goiter*. Report of the meeting of the PAHO Scientific Group on Research in Endemic Goiter held in Mexico, 1968 (J.B. Stanbury, ed.). PAHO Scientific Publication No. 193, Washington, 1969. pp. 378-380.

Fierro-Benítez, R.; Ramirez, I.; Estrella, E.; Jaramillo, C.; Diaz, C. and Urresta, J.: Iodized oil in the prevention of endemic goiter and associated defects in the Andean region of Ecuador. In: *Endemic Goiter*. Report of the meeting of the PAHO Scientific Group on Research in Endemic Goiter held in Mexico, 1968 (J.B. Stanbury, ed.). PAHO Scientific Publication No. 193, Washington, 1969. pp. 306-394.

Vernon, P.E.: *Intelligence and Cultural Environment*. Methuen, London, 1969.